March 20, 2008

The Honorable Susan Golding, Chair
MLPA Initiative Blue Ribbon Task Force
c/o California Resources Agency
1416 Ninth Street, Suite 1311
Sacramento, CA 95814

Dear Mayor Golding and fellow BRTF Members:

I am writing to identify modeling quality issues which I hope will be considered over the next month in SAT and BRTF meetings. The issues relate to understanding the proper usage and limitations of the SAT's population models, and quality assurance of the results that are presented to the BRTF and Fish and Game Commission. There is understandable enthusiasm for the excellent progress made in developing the models and the numerical scores that they produce. But, it is important to realize that given the amount of unknown habitat in the NCCS Region and the simplifications in creating the model cells, the potential to identify "clearly inferior" proposals is not possible for the NCCSR MPA evaluation. Yet, the initial display of results gives a first impression about packages, and I believe some erroneous impressions were set in the last round.

At the January 23 SAT meeting, the model results struck several RSG members as odd. We noted habitat gaps on the model maps, and wide differences when comparing similar MPA proposals. I raised my hand to volunteer to do a review and audit of the models to see if I could check the results. I have a background in software product development and data conversions and felt I would be able to do this. My review was limited in scope. My focus was on habitat percentages and a detail look at the habitat and MPA encoding for selected species and areas. This emphasis might be termed a review of data quality, but even in that category it was incomplete – I did not go back to source materials for example. Along the way, some bugs were found and fixed, but my effort should not be confused with adequate QA of the products. I had good cooperation with the UCD and Edom model teams and I am confident they responded directly to the areas under their control. I raised a number of issues in a series of reports and emails to the SAT, and my sense in conversations over the past week is that the broader issues have not yet been considered by the modeling committee. I may be wrong about that, but to facilitate responses, I recapped the main topics in the attached comment letter to the SAT. I think it is worth taking the time to have a good explanation of the proper usage of model results and QA of the data. One of the main findings in my audit was that the model results are highly dependent on the underlying habitat model which is a rough approximation of the actual habitat availability along the California Coast.

As a general rule it is best to release software products and documentation when one knows the product is ready rather than release the results to meet a deadline. In this case, the model consists not only of the software, but the data that describes the model cells. I urge you to support this effort and consider the SAT responses.

Sincerely,

Rick Johnson
NCCSR RSG Stakeholder (Alternate)

Attachment: letter to SAT on the same subject

March 20, 2008

Dr. Steve Morgan, Co-Chair
MLPA Initiative Science Advisory Team
c/o California Resources Agency
1416 Ninth Street, Suite 1311
Sacramento, CA 95814

Dear Dr. Morgan:

As I mentioned to you at the RSG meeting, I am concerned that there may be open items from my data quality audit reports on the population models. I believe these items require SAT attention. I covered many issues and details in those reports and want to highlight the specific topics that I hope the SAT will address in its next meeting, when round 3 model results are presented.

1. Proper usage and "Warning Label"
Whenever model results are presented in public, there should be a statement about proper usage, limitations and potential errors. .

Public presentations of the model results for round 2 showed rankings of proposals without any comment that the rankings may be inaccurate. In one SAT meeting some packages were described as clearly inferior and could be dropped. And, the first draft of the SAT's "synthesis of insights and results" report states: "The models also offer the potential to identify MPA configurations that are clearly inferior within the set of proposed packages." Given the amount of unknown habitat and the simplifications in creating the model cells and relationships, the potential to identify "clearly inferior" proposals is not possible in this NCCSR MPA process. It is correct that there is potential in the future development of the model, but it is misleading to me to suggest that the models have that capability now. When I ran the Edom model with a corrected habitat file, the list of "clearly inferior" proposals was reversed.

Carl Walters, the developer of the Edom model stated in an email: "I agree completely with his warning about not making so much of differences between plans that may depend on details of how the models represent those differences using aggregated cells."

In a discussion of the step of converting 2-dimensional GIS data into the 1-dimensional map used in the UCD model, Will White on the UCD Model team stated in an email: "I will admit that this method is fairly crude, but I would worry about trying to get too precise given the large stretches of nearshore habitat that are classified as 'unknown' in the MLPA-I's habitat maps, plus the fact that we don't have good habitat data for the region beyond state waters."

Both models are heavily weighted by low slope gradient areas where there are large areas mapped as hard bottom. According to fishermen, many of these areas are low relief and covered in sand (such as the wedge SE of Bodega Head, areas off ten mile beach and the section SE of Chimney Rock). In contrast, there are some areas where the deep water, 50 meter depth line runs close to shore such as off Point Reyes Headlands, around the Farallones, and off Black Point to Salt Point. In those locations with steep, high relief, presumably high quality habitat and ecotones, the model cells are missing the habitat because either it is coded as unknown or the 1 km square is too coarse to get the coding for a particular habitat type.

While discussing model limitations it should probably also be pointed out that there is scant information on local stock size, FLEP levels, steepness and recruitment compensation for nearshore fisheries. Estimates of these values drive the functioning of the models. One SAT member suspects weaker recruitment compensation than published NMFS/STARS assessments have assumed (i.e. greater sensitivity to over fishing), and another questioned the stock status and recruitment for Black Rockfish in California in the latest NMFS assessment.

I think the best use of the model results is to explore ideas, and raise questions. For example the difference between Proposal 0 and Proposal 4 probably illustrates something useful as the fishing effort varies. Some of the questions generated from modeling should be important for designing future monitoring and research. Any suggestion that the model results can replace the SAT size and spacing guidelines seems speculative to me.

2. Farallones statement of limitations

Information about the models should state the difficulty representing the Farallones in the models. My understanding is there is no solid information about how to model the link between the Farallones and coastal populations. So, the Farallones were added to the south end of the models instead of inserting them in the middle between Pt Reyes and Moss Beach. Also, because of the way the 1 dimensional model cells are developed and the way MPA impacts are coded as horizontal strips across the Farallones, I estimate the Farallon contribution to total yield in the models is about 1/3 what might be expected just on the amount of habitat in the subregion.

3. Habitat files - What habitat files will be used for round 3 model runs?

My audit reports demonstrated the importance of the habitat files in determining model results. Much of the source of difficulty is the 30% of shallow water habitat that is classified as "unknown" in the NCCSR GIS data layers.

The UCD model codes the presence or absence of habitat in each 1 km model cell. The bias related to excluding unknown shallow habitat is much reduced for UCD because UCD modified their procedure to estimate many of the unknown 1 km square blocks. One could argue the remaining unknowns are insignificant for UCD

given its model design. But, UCD does miss habitat presence in the areas where the 50 meter depth runs close to shore. I suggested changing the classification sensitivity in those areas.

What habitat file will be used for round 3 of the Edom model? Edom is even more sensitive to habitat coding errors because the model uses % habitat type in a model cell. I built a corrected habitat file which I think is OK, but that needs to be confirmed and accepted or a new habitat file must be created. I suggested some ways to do that, but I do not know the status. If a new habitat file is created there are many adjustments needed for unknown habitat, estuaries and handling the reef and bend around Duxbury. Quality assurance is needed if the Edom habitat data file is changed.

I submitted Kvitek maps showing rock habitat along the Pt Reyes headlands and ask that that be coded in the models. I did code it in the corrected file which I made for Edom.

As a side note: I am actually surprised that the model process takes the GIS data which is continuous, breaks it into 1 km squares, and then groups it back into 1km model strips to compute habitat percentages. I think if someone takes the time to define the 242 model cells as shapes, then you can extract whatever data you want from the GIS system and put it into model cells.

4. MPA coding – who will do QA of MPA coding for the models?
Coding the MPA coverage in model cells is error prone. MPA coding will be done for round 3 on a tight time line, and we can be certain that mistakes will be made. I don't think time has been allowed for QA. Some provision is needed to check the MPA coding. I believe staff time is needed for this as I do not have time to go through all the MPAs for two models.

There are also some issues around MPA coding. What percent presence of an MPA in a model cell counts as protection? Edom uses 10% which seems too low a threshold for complete protection. I did not get the threshold information for the UCD model. There may be some inadvertent gaming of the models by designing thin wedges of MPA boundaries that cross many cells but provide little protection (south of Bodega, south of Pillar Point and along west side of Drakes Bay where the model cells are north-south.)

The UCD model counts state marine parks as fractional protection of species which seems overly optimistic to me.

Edom does not use species specific MPA designation – crab allowance is frequent and one or two SMCA mention halibut take allowed.

5. What benchmarks, if any, will be done for the NCCSR project and when?
Dave Rogers, a mechanical engineer who is a member of Coastside Fishing Club, made comments at the last SAT meeting, which according to my notes were:

> You should benchmark the models to real world conditions. If the model outputs predict what is going on now, then you can be more confident that the model works and use it to predict results with MPAs. I challenge you to do benchmarks.

Subsequently, I outlined some ideas for benchmarks. The concept is to predict something using the models that can actually be confirmed with known observations. That would give some confidence that the models will help predict new scenarios that cannot be confirmed. Here are a few illustrative examples, but I expect the SAT could do a better job than me defining the appropriate benchmarks:

> 1. Predict where the fish biomass of catch is distributed proportionately along the coast, in the absence of MPAs. If there are distributions of historical fishery catch data, that information can be a benchmark to confirm the model outputs. If data is only available for certain subregions studying those subregions could be better than no benchmark at all.
>
> 2. Predict where the fishermen go, in the absence of MPAs. The Ecotrust data could be aggregated by 1km strips matching the model cells, and then see if the models predict where the fishermen would go. (Ecotrust comparison could be done for crab, halibut and rockfish). The differences might give a sense of error ranges in model results, too.
>
> 3. Test the stock recruitment and vulnerability to catch by age assumptions by predicting the differential effects of fishing in a mixed species fishery, in the absence of MPAs and with MPAs. I was interested in this one because it deals with an area where MPAS should be helpful. I picked the Copper Rockfish because it is a late maturing fish, which is disproportionately depleted by fishing effort, and seems to be disappearing from the fishery. Copper Rockfish are part of the nearshore fishery group. There is evidence from Point Lobos studies that an MPA protects Copper Rockfish. I prepared a simple model with Edom that showed a fish could be nearly eliminated from the mixed fishery without MPAs and survives with an MPA network.

6. Audit packet
For consumers of the information and anyone who wants to check the data in the future, what would help is to have an audit packet for each round of MLPA analysis including:

Habitat 2 D grid including unknown cells and land cells so that the grid can be lined up with a map. (or as mentioned earlier, define model cell shapes in the GIS and convert directly from GIS to 1D model cells.)
1D Model Cell boundary reference – an index of the north latitude line for each cell (DD MM.MM) which will allow one to link the model cells to a map. (and west longitude line where the model goes vertical)
2D to 1D cross reference spreadsheet if available to help to find coding errors.
1D habitat spreadsheet (UCD absence/presence of habitat, and Edom percentage habitat by type). This spreadsheet can be checked for accuracy by cell and the subregion totals can be compared to profile values to see if the model is reasonably close to values calculated from the GIS system.
Once the habitat data are set and QA'ed for a region they will not change through each draft MPA round.

For each MPA proposal round:
2D MPA grid coding of MPA coverage if available
1D MPA spreadsheet showing coding by model cell of MPA protection level by species (Edom has only one value, UCD carries MPA coverage by species and fractional coverage levels)
Once the MPA data are set and QA'ed for an MPA proposal round they will not change through each model run.

For each model run. It would include:
Input settings: e.g. mix of species and their weighting in the total, fishing effort and type of effort model (this information should allow another user to replicate the run)
Output table: values by species of biomass, catch, economic value, FLEP by model cell
(This information is not needed for every run, but whenever results are being checked in detail or compared to benchmark values, the detail is necessary.)

With all of this information one can check the data quality of the models, and examine the results for reasonableness.

7. What species mix will be used for overall sustainability graphs?
In the BRTF and RSG presentations of round 2 model results, many species were mixed and weighted to create a single result provided for economic and sustainability measures. It is helpful to know the species mix and weighting underlying a graph that is displayed. The aggregated all-species view may be useful, but I think displays aggregated for the nearshore fishery management species would also be valuable since the nearshore fishery management species are most likely to be affected by MPAs. Putting in crab and halibut may give an overall sense of MPA effects, but it clouds the picture since there is so much uncertainty about MPA effects on those species. I don't know which way is best to use the model but I ask the SAT to consider showing both.

I will appreciate consideration of these topics by the SAT and responses at the SAT meeting. It is probably appropriate to have a quality assurance report on the habitat and MPA data used in the model runs for the evaluations. Thank you.

Sincerely,

Rick Johnson
NCCSR RSG Stakeholder (Alternate)